



OIPPE

ENTERED

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/717,450

DATE: 01/24/2003

TIME: 12:50:13

#10

Input Set : A:\Sequence Listing (ASCII copy).txt

Output Set: N:\CRF4\01242003\I717450.raw

## SEQUENCE LISTING

## C--&gt; 5 (1) GENERAL INFORMATION:

7 (i) APPLICANT: Neuhold, Lisa

8 Killar, Loran

## C--&gt; 10 (ii) TITLE OF INVENTION: TRANSGENIC ANIMAL MODEL FOR

11 DEGENERATIVE DISEASES OF CARTILAGE

13 (iii) NUMBER OF SEQUENCES: 21

15 (iv) CORRESPONDENCE ADDRESS:

16 (A) ADDRESSEE: Darby &amp; Darby PC

17 (B) STREET: 805 Third Avenue

18 (C) CITY: New York

19 (D) STATE: NY

20 (E) COUNTRY: USA

21 (F) ZIP: 10022

23 (v) COMPUTER READABLE FORM:

24 (A) MEDIUM TYPE: Diskette

25 (B) COMPUTER: IBM Compatible

26 (C) OPERATING SYSTEM: DOS

27 (D) SOFTWARE: FastSEQ for Windows Version 2.0

29 (vi) CURRENT APPLICATION DATA:

C--&gt; 30 (A) APPLICATION NUMBER: US/09/717,450

C--&gt; 31 (B) FILING DATE: 10-Jan-2003

32 (C) CLASSIFICATION:

34 (vii) PRIOR APPLICATION DATA:

35 (A) APPLICATION NUMBER:

36 (B) FILING DATE:

38 (viii) ATTORNEY/AGENT INFORMATION:

39 (A) NAME: Green, Reza

40 (B) REGISTRATION NUMBER: 38,475

41 (C) REFERENCE/DOCKET NUMBER: 0630/0D532

43 (ix) TELECOMMUNICATION INFORMATION:

44 (A) TELEPHONE: 212-527-7700

45 (B) TELEFAX: 212-753-6237

46 (C) TELEX: 236687

49 (2) INFORMATION FOR SEQ ID NO: 1:

51 (i) SEQUENCE CHARACTERISTICS:

52 (A) LENGTH: 471 amino acids

53 (B) TYPE: amino acid

54 (C) STRANDEDNESS: single

55 (D) TOPOLOGY: linear

57 (ii) MOLECULE TYPE: protein

59 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

61 Met His Pro Gly Val Leu Ala Ala Phe Leu Phe Leu Ser Trp Thr His

RECEIVED

FEB 20 2003

TECH CENTER 1600/2900

## RAW SEQUENCE LISTING

DATE: 01/24/2003

PATENT APPLICATION: US/09/717,450

TIME: 12:50:13

Input Set : A:\Sequence Listing (ASCII copy).txt

Output Set: N:\CRF4\01242003\I717450.raw

```

62      1              5              10              15
63 Cys Arg Ala Leu Pro Leu Pro Ser Gly Gly Asp Glu Asp Asp Leu Ser
64              20              25              30
65 Glu Glu Asp Leu Gln Phe Ala Glu Arg Tyr Leu Arg Ser Tyr Tyr His
66              35              40              45
67 Pro Thr Asn Leu Ala Gly Ile Leu Lys Glu Asn Ala Ala Ser Ser Met
68              50              55              60
69 Thr Glu Arg Leu Arg Glu Met Gln Ser Phe Phe Gly Leu Glu Val Thr
70              65              70              75              80
71 Gly Lys Leu Asp Asp Asn Thr Leu Asp Val Met Lys Lys Pro Arg Cys
72              85              90              95
73 Gly Val Val Asp Val Gly Glu Tyr Asn Val Phe Pro Arg Thr Leu Lys
74              100              105              110
75 Trp Ser Lys Met Asn Leu Thr Tyr Arg Ile Val Asn Tyr Thr Pro Asp
76              115              120              125
77 Met Thr His Ser Glu Val Glu Lys Ala Phe Lys Lys Ala Phe Lys Val
78              130              135              140
79 Trp Ser Asp Val Thr Pro Leu Asn Phe Thr Arg Leu His Asp Gly Ile
80              145              150              155              160
81 Ala Asp Ile Met Ile Ser Phe Gly Ile Lys Glu His Gly Asp Phe Tyr
82              165              170              175
83 Pro Phe Asp Gly Pro Ser Gly Leu Leu Ala His Ala Phe Pro Pro Gly
84              180              185              190
85 Pro Asn Tyr Gly Gly Asp Ala His Phe Asp Asp Asp Glu Thr Trp Thr
86              195              200              205
87 Ser Ser Ser Lys Gly Tyr Asn Leu Phe Leu Val Ala Ala His Glu Phe
88              210              215              220
89 Gly His Ser Leu Gly Leu Asp His Ser Lys Asp Pro Gly Ala Leu Met
90              225              230              235              240
91 Phe Pro Ile Tyr Thr Tyr Thr Gly Lys Ser His Phe Met Leu Pro Asp
92              245              250              255
93 Asp Asp Val Gln Gly Ile Gln Ser Leu Tyr Gly Pro Gly Asp Glu Asp
94              260              265              270
95 Pro Asn Pro Lys His Pro Lys Thr Pro Asp Lys Cys Asp Pro Ser Leu
96              275              280              285
97 Ser Leu Asp Ala Ile Thr Ser Leu Arg Gly Glu Thr Met Ile Phe Lys
98              290              295              300
99 Asp Arg Phe Phe Trp Arg Leu His Pro Gln Gln Val Asp Ala Glu Leu
100              305              310              315              320
101 Phe Leu Thr Lys Ser Phe Trp Pro Glu Leu Pro Asn Arg Ile Asp Ala
102              325              330              335
103 Ala Tyr Glu His Pro Ser His Asp Leu Ile Phe Ile Phe Arg Gly Arg
104              340              345              350
105 Lys Phe Trp Ala Leu Asn Gly Tyr Asp Ile Leu Glu Gly Tyr Pro Lys
106              355              360              365
107 Lys Ile Ser Glu Leu Gly Leu Pro Lys Glu Val Lys Lys Ile Ser Ala
108              370              375              380
109 Ala Val His Phe Glu Asp Thr Gly Lys Thr Leu Leu Phe Ser Gly Asn
110              385              390              395              400

```

## RAW SEQUENCE LISTING

DATE: 01/24/2003

PATENT APPLICATION: US/09/717,450

TIME: 12:50:13

Input Set : A:\Sequence Listing (ASCII copy).txt

Output Set: N:\CRF4\01242003\I717450.raw

```

111 Gln Val Trp Arg Tyr Asp Asp Thr Asn His Ile Met Asp Lys Asp Tyr
112                               405                               410                               415
113 Pro Arg Leu Ile Glu Glu Asp Phe Pro Gly Ile Gly Asp Lys Val Asp
114                               420                               425                               430
115 Ala Val Tyr Glu Lys Asn Gly Tyr Ile Tyr Phe Phe Asn Gly Pro Ile
116                               435                               440                               445
117 Gln Phe Glu Tyr Ser Ile Trp Ser Asn Arg Ile Val Arg Val Met Pro
118                               450                               455                               460
119 Ala Asn Ser Ile Leu Trp Cys
120 465                               470
122 (2) INFORMATION FOR SEQ ID NO: 2:
124 (i) SEQUENCE CHARACTERISTICS:
125 (A) LENGTH: 470 base pairs
126 (B) TYPE: nucleic acid
127 (C) STRANDEDNESS: single
128 (D) TOPOLOGY: linear
131 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
133 CTCGAGTTTA CCACTCCCTA TCAGTGATAG AGAAAAGTGA AAGTCGAGTT TACCACTCCC      60
134 TATCAGTGAT AGAGAAAAGT GAAAGTCGAG TTTACCACTC CCTATCAGTG ATAGAGAAAA      120
135 GTGAAAGTCG AGTTTACCAC TCCCTATCAG TGATAGAGAA AAGTGAAAGT CGAGTTTACC      180
136 ACTCCCTATC AGTGATAGAG AAAAGTGAAA GTCGAGTTTA CCACTCCCTA TCAGTGATAG      240
137 AGAAAAGTGA AAGTCGAGTT TACCACTCCC TATCAGTGAT AGAGAAAAGT GAAAGTCGAG      300
138 CTCGGTACCC GGGTCGAGTA GCGGTGTACG GTGGGAGGCC TATATAAGCA GAGCTCGTTT      360
139 AGTGAACCGT CAGATCGCCT GGAGACGCCA TCCACGCTGT TTTGACCTCC ATAGAAGACA      420
140 CCGGGACCGA TCCAGCCTCC GCGGCCCCGA ATTAGCTTGA TATCGAATTC      470
142 (2) INFORMATION FOR SEQ ID NO: 3:
144 (i) SEQUENCE CHARACTERISTICS:
145 (A) LENGTH: 3479 base pairs
146 (B) TYPE: nucleic acid
147 (C) STRANDEDNESS: single
148 (D) TOPOLOGY: linear
151 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
153 GGTACCACTA GTAAGCTTAG ATCCACTGTC TGGGATTATA TCAGGACAAC CGAAGCCTGG      60
154 AAAGTGTATT AGGTAGAGCA TTTTCTTCCA CGTGTGTTGG CACGTTTCCG ACAGCTAGGA      120
155 TTCCAGCTCT GTCTTTGTAT GTTACAGACT GTAAATCAAT CGCAGGTGAA ACTGTTTGGA      180
156 CAGTAGGTGG GGATCAAAGA CCCTCCGCCC GTGAGACTCT AGGCGCTTTC CCCTGCCACC      240
157 AGCCTGTCTC CAGAGATGCT CTGGAAGGAG GCGGGCCCCG GCGGTCTTTC TGCTCTTTAG      300
158 CGTGGCGGAC GCGGCGGCGG GGGCAGGGCT GGAGCAGAGA GCGCTGCAGT GATAGAACTT      360
159 TCTGACCCCG CTGCGCAGGG CGGCAGGGTG GCAGGGTGGC AGGGTGGCGA GCTAAGCCAG      420
160 AGCCGAACGC TGGAGCTCTG GGAGGAACAT CGAAGGTTTG TATGTGGTCT GAGATCGGCC      480
161 TGAATATATT TTTTGTCTCT AAATTTGCAA GCACACACCC ACAAAGCTGC GGTCTTGACC      540
162 GGTATTCTTT ATAGAGCGCA ATGGAGTGAG CTGAGTGTCT AAACGATTTC CCTAATTCAT      600
163 CTGATAGCAG AGGCGCTCTC CTAATTGGCG AAGAGCTGCC TCATGTCCGC AACTTTTTTG      660
164 CAGAGTGAAT TCCACAGCTT TGTGTGTGTG TGTGGGGGGG GGTGTAAGGG GTGTCTAAAA      720
165 CTTTCGGTCT CCTACTATTC TGTATCTCGA CCGGTTGGTT TTACACCCCG GCTCATCTCA      780
166 TCAACGCAAA CACCCCCACT CTCCTATGGA CCAAGGACC TGACGTGGGG GAAGGTGGAC      840
167 ATTAGGAATG TCAGAAACCT AGAGTCCACG CTCCTCCTCT CCATCTTTCC ACGAGTTTGG      900
168 GAAACTTCTT GGCTGCGAAG ACTTTGACCC ACATCTGCAT TTCTCAGCCC CAGCTTCCAA      960
169 AAGTGCTGCA GGTTCGGGAG GGGAGACCTC AGTCCTCCTT TGTGAGGCTT GTTTGCGTTG      1020

```

## RAW SEQUENCE LISTING

DATE: 01/24/2003

PATENT APPLICATION: US/09/717,450

TIME: 12:50:13

Input Set : A:\Sequence Listing (ASCII copy).txt

Output Set: N:\CRF4\01242003\I717450.raw

```

170 AGGGATTGGC AGCGATGGCT TCCAGATGGG CTGAAACCCT GCCCGTATTT ATTTAAACTG 1080
171 GTTCCTCGTG GAGAGCTGTG AATCGGGCTC TGTATGCGCT CGAGAAAAGC CCCATTCATG 1140
172 AGAGGCAAGG CCCAGTGGGT CCCCCGACT CCCCACCC CTTCTCCAC AATATATCCC 1200
173 CCCTCCCTGT GCCCGCCTGC CGCCACCTCC CGGGCTCCGG CCCCAGCGC AGCGGCGACG 1260
174 AAGCAACACA GTTCCCCGAA AGAGGTAGCT TTTTAATTGG CCAGCCACAA AGAATCACTT 1320
175 ATGCCGCACG GCGGTAACGA GGGGAACCGG ATCGGGCGGC CAGGATGCTA TCTGTGTAGC 1380
176 CCTTTTCGTG CCACAATTAG GGTGGTGTG GCTTCCTCCG ACCGCACCTA GGCGATCTGG 1440
177 TTACACTGTT GGCTCCTTTC TTGGGCAGTC ATTTAATCCT ACTTTTTACT CTACGAATGT 1500
178 CTGTCTGATG GAGGGCTGTG TCCGGAGCCC CATCCACAAA GAGTCAGCCA GCAGCTCTCA 1560
179 CACCCGGCTG GATCTCATAT GGTGCACTCT CAGTACAATC TGCTCTGATG CCGCATAGTT 1620
180 AAGCCAGCCA AGCTAGCTTG CGCAAGCTAG CTTGCGATCC GTAAAAATGT GTGAGAGTTA 1680
181 CAAAATGTCT TCCGGGCTAA GATCCGACAG CCATGGTCCA AAGAAGACTT CGGCACTGCA 1740
182 GACTTAAAC CAGCTTTCTA GCAGAGGCAG AAGGATCTAG AGCCAAAGGC AAAGACTTGA 1800
183 ATAGGCTGGG AAGATGCAAG AATGGCATT TACATAAAGA AACTCTCTC CTTTTCCAGC 1860
184 CAGCACACTT GCATAGAAAT TAAGTTTTAC ACTTGAAGTT CTTTGTTC ATCCTGAGAA 1920
185 GCTCCAAAGT CTGAGGTGGT GTGGTATGCT GGGTAATTCT CCCCACCCC CAACATTCCC 1980
186 TGGGGGTTCC ATGGGGGTAG CTTCTCCAA GGAATTCCAG CGGCAACACA GAAATCCCAC 2040
187 TTCGAGACAA AGGAGTTACT GCTTAAATCA GGCCCTAATT TCCAAGGTT CTTTGCTTA 2100
188 AAGTTCCCTA GAGGACCATC TCACTTCTAA AGAAAAGGTG TATTCGGGGA CCCATCCTCA 2160
189 ACCTCCTTGT TATGGAAGGA GACTTCGGGA ACAGAGCAAG GGCTGAGCCT CCGGCAGTTT 2220
190 GGGGTAAGGT TGGGGTTGGG GGGAGCAAGG AAGGCAAGTG AGGCTGGAGG CCCAGGGATA 2280
191 GGGGAAGATG TGTGTGTGTG TGTGTGTGTG TGTGTGTGTG TGTGTGTGTC TCGGGGATGG 2340
192 TGGTGGTGGA CAACTAGGAA ACTCTGGCGC TTTCTCCTCC CTCACAAAA CTGAGTCCAG 2400
193 CTGGAGCCGC CTCAGACTC TCTGGCCAG GCCTCAGAGT GGTCAACAGT CCCTGGCCAG 2460
194 CGTTGCTCTC TCCAGGCTAA GGGCACCCAG TCCCCTGGAG ATTCTGAAC CTGGGCCAGG 2520
195 AAGAGCCGAA TTAGACAAGT GTCTCCAATC CGGCTGCGTG CGGATTTTGT TGCGGTGTCC 2580
196 CTCGGTTGTC TGCAGTTTCT TTAGTCCCTT CCCTGGCCTG CCCCTTACAC CTCCACACAG 2640
197 GTCCCCCTCT GTGTAGGAAT ACACCAGACC CTCTCTTAGC CACACACACC TCCAGTCCCC 2700
198 CGTCTACCTA GATTTTTTTC ATAGCTAGTT GGATGGGGGA TGGGTTAGGG AGGCTGGGTT 2760
199 TGCGAGCCTC CAGGTGGGAG TTCACGACA GGTACTCCGC AAAGGAGCTG GAAGGCAGGT 2820
200 CTGGAAAAC GTCCCCAGA TTTAGGATTC TGGGCAGCTT CCATCAGCTT ATACTTTGGC 2880
201 TCCCCCGCCC CCTAACTCC CCATCCCCAC CTTCTTTCT CCCGTTACTT CGTCCTCCCT 2940
202 CGCCTTTCCA GCCTTGAGTC TAAAGCTCCA TGCTTATGCC TCTGCAAACA ACCCCCTCCC 3000
203 TTCTAACCCC AGCAGAACTC CGAGGAAAGG GGCCGGAGGC CCCCCTTCTC GCCTGTGGTT 3060
204 AGAGGGGGCA GTGTGGCAGT CCCAAGTGGG GGCGACCGGA GGCCGTCTCG GTGCCCCGCC 3120
205 CGATCAGGCC ACTGGGCACA TCGGGGGCGG GAAGCTGGGC TCACCAAAGG GGCGACTGGC 3180
206 CTTGGCAGGT GTGGGCTCTG GTCCGGCCTG GGCAGGCTCC GGGGGCGGGG TCTCAGGTTA 3240
207 CAGCCCCGCG GGGGGCTGGG GGGCGGCCCG CGGTTTGGGC TGGTTTGCCA GCCTTTGGAG 3300
208 CGACCGGGAG CATATAACCG GAGCCTCTGC TGGGAGAAGA CGCAGAGCGC CGCTGGGCTG 3360
209 CCGGGTCTCC TGCCTCCTCC TCCTGCTCCT AGAGCCTCCT GCATGAGGGC GCGGTAGAGA 3420
210 CCCGGACCCG CTCCGTGCTC TGCCGCCTCG CCGAGCTTCG CCCGCAAGCT GGGGAATTC 3479

```

212 (2) INFORMATION FOR SEQ ID NO: 4:

214 (i) SEQUENCE CHARACTERISTICS:

215 (A) LENGTH: 8 amino acids

216 (B) TYPE: amino acid

217 (C) STRANDEDNESS: single

218 (D) TOPOLOGY: linear

220 (ii) MOLECULE TYPE: peptide

222 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

## RAW SEQUENCE LISTING

DATE: 01/24/2003

PATENT APPLICATION: US/09/717,450

TIME: 12:50:13

Input Set : A:\Sequence Listing (ASCII copy).txt

Output Set: N:\CRF4\01242003\I717450.raw

```

224 Pro Arg Cys Gly Val Pro Asp Val
225 1 5
227 (2) INFORMATION FOR SEQ ID NO: 5:
229 (i) SEQUENCE CHARACTERISTICS:
230 (A) LENGTH: 39 base pairs
231 (B) TYPE: nucleic acid
232 (C) STRANDEDNESS: single
233 (D) TOPOLOGY: linear
236 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
238 AAGCCAAGAT GCGGGGTTGT CGATGTGGGT GAATACAAT 39
240 (2) INFORMATION FOR SEQ ID NO: 6:
242 (i) SEQUENCE CHARACTERISTICS:
243 (A) LENGTH: 40 base pairs
244 (B) TYPE: nucleic acid
245 (C) STRANDEDNESS: single
246 (D) TOPOLOGY: linear
249 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
251 GAAAAAGCCA AGATGCGGGG GTCCTGATGT GGGTGAATAC 40
253 (2) INFORMATION FOR SEQ ID NO: 7:
255 (i) SEQUENCE CHARACTERISTICS:
256 (A) LENGTH: 98 base pairs
257 (B) TYPE: nucleic acid
258 (C) STRANDEDNESS: single
259 (D) TOPOLOGY: linear
262 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
264 GGTACCACTA GTAAGCTTAG ATCTCATATG GTCGACCCCG GGGAATTCCT GCAGGGATCC 60
265 TCTAGAAGTA CTCCATGGGT ATACATCGAT GCGGCCGC 98
267 (2) INFORMATION FOR SEQ ID NO: 8:
269 (i) SEQUENCE CHARACTERISTICS:
270 (A) LENGTH: 2792 base pairs
271 (B) TYPE: nucleic acid
272 (C) STRANDEDNESS: single
273 (D) TOPOLOGY: linear
275 (ii) MOLECULE TYPE: cDNA
277 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:
279 CTCGAGTTTA CCACTCCCTA TCAGTGATAG AGAAAAGTGA AAGTCGAGTT TACCACTCCC 60
280 TATCAGTGAT AGAGAAAAGT GAAAGTCGAG TTTACCACTC CCTATCAGTG ATAGAGAAAA 120
281 GTGAAAGTCG AGTTTACCAC TCCCTATCAG TGATAGAGAA AAGTGAAAGT CGAGTTTACC 180
282 ACTCCCTATC AGTGATAGAG AAAAGTGAAA GTCGAGTTTA CCACTCCCTA TCAGTGATAG 240
283 AGAAAAGTGA AAGTCGAGTT TACCACTCCC TATCAGTGAT AGAGAAAAGT GAAAGTCGAG 300
284 CTCGGTACCC GGGTCGAGTA GCGGTGTACG GTGGGAGGCC TATATAAGCA GAGCTCGTTT 360
285 AGTGAACCGT CAGATCGCCT GGAGACGCCA TCCACGCTGT TTTGACCTCC ATAGAAGACA 420
286 CCGGGACCGA TCCAGCCTCC GCGGCCCCGA ATTAGCTTGA TATCGAATTC GAGCTCGGTA 480
287 CCCGGGGATC CTCTAGACAA GATGCATCCA GGGGTCCTGG CTGCCTTCCT CTTCTTGAGC 540
288 TGGACTCATT GTCGGGCCCT GCCCCTTCCC AGTGGTGGTG ATGAAGATGA TTTGTCTGAG 600
289 GAAGACCTCC AGTTTGCAGA GCGCTACCTG AGATCATACT ACCATCCTAC AAATCTCGCG 660
290 GGAATCCTGA AGGAGAATGC AGCAAGCTCC ATGACTGAGA GGCTCCGAGA AATGCAGTCT 720
291 TTCTTCGGCT TAGAGGTGAC TGGCAAACCT GACGATAACA CCTTAGATGT CATGAAAAAG 780
292 CCAAGATGCG GGGTTGTCGA TGTGGGTGAA TACAATGTTT TCCCTCGAAC TCTTAAATGG 840

```

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/09/717,450

DATE: 01/24/2003  
TIME: 12:50:14

Input Set : A:\Sequence Listing (ASCII copy).txt  
Output Set: N:\CRF4\01242003\I717450.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:19; Xaa Pos.5

Seq#:20; Xaa Pos.3,6,7,8,9,10

**VERIFICATION SUMMARY**

DATE: 01/24/2003

PATENT APPLICATION: US/09/717,450

TIME: 12:50:14

Input Set : A:\Sequence Listing (ASCII copy).txt

Output Set: N:\CRF4\01242003\I717450.raw

L:5 M:220 C: Keyword misspelled or invalid format, [(1) GENERAL INFORMATION:]  
L:10 M:220 C: Keyword misspelled or invalid format, [(ii) TITLE OF INVENTION:]  
L:30 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]  
L:31 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]  
L:715 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:19 after pos.:0  
L:731 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20 after pos.:0